Study of T7 RNA Polymerase-DNA interactions by way of Unnatural Base Pairs

Brian Vial
Spring 2010
5/12/2010

In partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in the School of Biological Sciences

_____ _____

(Bradley Hall) Date
Supervising Professor

_____ _____

(Andrew Ellington, PhD) Date
Co-Supervising Professor

_____ _____

(Ruth Buskirk, PhD) Date
Honors Advisor in School of Biological Sciences

Abstract

The use of the T7 RNA polymerase is ubiquitous in molecular biology as a research tool for the study of the regulation and mechanism of transcription in biological systems. Previous research has demonstrated its high promoter specificity; therefore, it is commonly used for transcription of DNA cloned into bacterial vectors[1]. This high promoter specificity also ensures that any modified or mutated promoter likely will cause the T7 RNA polymerase to bind inefficiently when encountering an altered base and cease the transcription of its DNA substrate. Especially critical are the base pairs at the -16 to +6 positions relative to the start site. Many of these base pairs are involved in the initial binding of the T7 RNA polymerase protein and the formation of the transcription bubble[1]. This experiment focused on the effects of altering these critical base pairs upstream of the transcription start site in conjunction with the modification of the T7 RNA polymerase to gain a better understanding of protein-DNA interactions and yield possible new techniques in research[2].

Specifically, the major groove interactions between the T7 RNA polymerase and bases in the -9 to -7 position relative to the start site were studied[1]. Wild-type bases were replaced with a combination of wild-type substitutions, iso-forms of wild-type bases, and artificial bases with modified functional groups. In this experiment, both iso-Guanine and iso-Cytosine pairs were substituted at multiple locations in the promoter. Furthermore, a completely unnatural base pair of dZ:dP was used for substitutions. These are unique due to the dP base missing an N7 group and an NO_2 being added to the dZ base.

In addition to these promoter alterations, multiple variants of the wild-type T7 RNA polymerase were produced by replacing amino acids responsible for the protein-DNA interactions seen in transcription. These variants were tested with modified promoters to

ascertain the success of transcription. Within the time constraints of this project, moderate success was found when some variants increased the transcription activity normally seen between the modified promoters and the wild-type T7 RNA polymerase.

Future studies in this field may include testing more variant T7 RNA polymerases and decoding details of DNA-protein interactions. Furthermore, this experiment extends the possibility of obtaining a modified promoter with a compatible T7 RNA Polymerase where both must concurrently be present for transcription to successfully occur. With this new research technique, single genes could be up-regulated or down-regulated with a high degree of accuracy due to the uniqueness of the promoter sequence. An unnatural promoter sequence with little tolerance of wild-type polymerases would ensure that no other genes are upregulated when using the matching modified T7 RNA polymerase.

Chapter 1: Background of DNA-protein interaction research

Introduction

Unnatural base pairs previously have been used to study important DNA-protein interactions such as how DNA polymerase recognizes base-pairing in DNA and leads to an increase in effective replication[2]. The early Watson-Crick studies of the DNA double helix structure determined that base pairing in DNA follows two basic rules, base "*complementarity"* and "*hydrogen bond complemetarity*[5]." Those studies entailed the basic principle that pyrimidines and purines with similar hydrogen bonding will always pair together and provide DNA its stable structure. However, later studies showed that DNA actually may maintain stable structures when unnatural base pairs are substituted, even with those of different sizes[6].

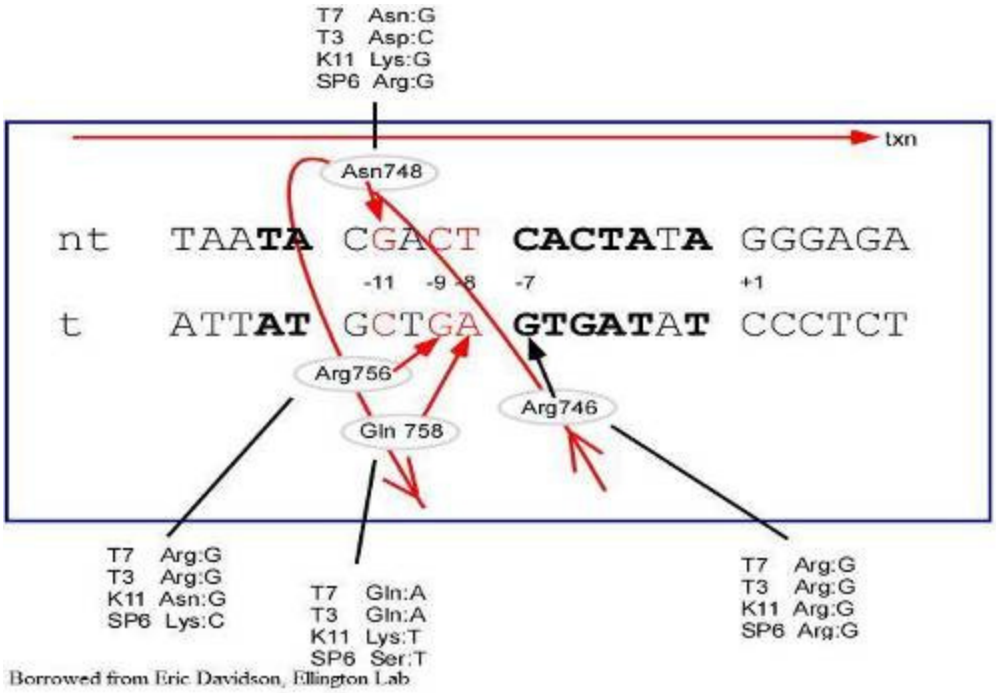
Furthermore, the concept of unnatural base pair incorporation in DNA gave science new avenues to explore in DNA replication. One study showed that even when unnatural base pairs are incorporated into DNA, the T7 RNA polymerase remains able to amplify RNA transcription and incorporate the unnatural base analogs into new RNA molecules[7]. These results have important implications for RNA substrates. Large RNA molecules that normally serve as reaction components or ribozymes may have their basic functional structure altered by the incorporation of unnatural base pairs with modified functional groups. This discovery may lead to an array of RNA applications in synthetic biology in which alternate functional groups are added to the modified RNA molecule to fine tune its purpose and abilities inside living systems.

Finally, unnatural base pair incorporation into DNA leads to more diversity in the genetic code that may have a variety of applications in synthetic biology. If unnatural base pairs are found to function well in the transcription/translation machinery *in vivo*, the normal breadth of translation processes may be expanded. One experiment demonstrated that tRNAs with unnatural bases incorporated in their anti-codons could be charged by a variant tRNA synthetase while the wild-type tRNA synthetase did not recognize the unnatural tRNAs[3]. These findings suggest the abilities of unnatural bases to alter translation processes but also illustrate their ability to function in normal translation machinery when exposed to the right conditions, in this case the variant tRNA synthetase. For synthetic biology purposes, this discovery could lead to an expansion in the number of codons available in the translation process. In addition, if variant tRNAs can be charged with modified amino acids, the ability to incorporate variant amino acids into newly synthesized proteins may be likely, and this may lead to the engineering and consequent study of interesting and more efficient proteins.

Chapter 2: Experimental findings on unnatural base pair substitutions and their effect on the T7 RNA Polymerase

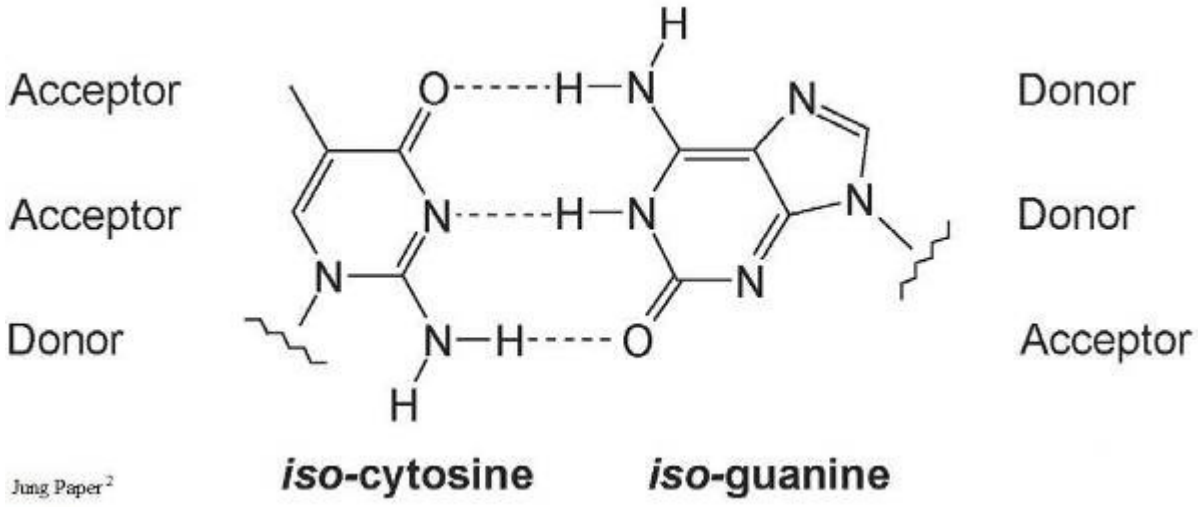
Background

Previous research has shown that the success of T7 RNA polymerase depends heavily on a highly conserved promoter sequence located at the -16 to +6 positions relative to the start site[1]. There are four critical base interactions that occur between the T7 RNA polymerase and the promoter. The Asn748 side-chain binds with the guanine located at the -11 position and is the only interaction that occurs on the non-template strand. The other three critical interactions occur at the -9 through -7 positions. Here Arg756, Gln758, and Arg746 bind with -9-Guanine, -8-Adenine, and -7-Guanine, respectively. **Figure 1** illustrates this major interaction between the amino acid groups of T7 and their corresponding base. In addition, all four interactions occur on the major groove[1].



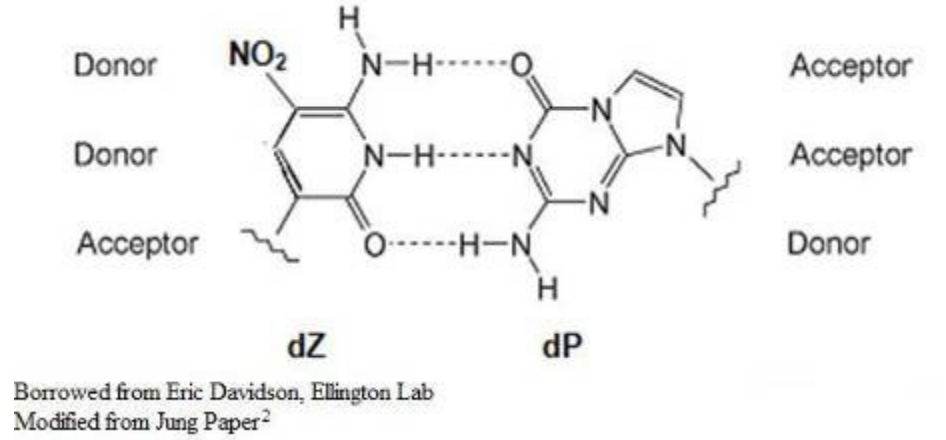
(Figure 1) Major groove interactions at the -11, -9, -8, and -7 positions between promoter and amino acid residues in the T7 RNA Polymerase

Due to the importance of these four interactions, these bases were targeted for substitution with unnatural ones. The major substitutions performed involved iso-Cytosine:iso-Guanine and dZ:dP base pairing. The iso-C:iso-G pairing is interesting because of its similarity to the orientation of adenine:thymine base pairing with respect to the major groove (See **Figure 2**). Study of these provided insight into the issue of whether base recognition or major groove interactions are the most important factor for determining successful T7 RNA polymerase binding and transcription.



(**Figure 2**) Iso-cytosine and iso-guanine base pairing when incorporated into DNA

On the other hand, studying the interactions between the dZ:dP base pairing and the T7 RNA polymerase provided additional information about factors involved in these major groove interactions. The dZ:dP base pairs are unique due to the nitrogen dioxide group located on the dZ base (**Figure 3**). A completely unnatural functional group such as nitrogen dioxide provides variability in the genetic code and may lead to new applications in synthetic biology.



(**Figure 3**) dZ:dP base pairing when incorporated into DNA

Overall, with 12 possible substitutions with unnatural base pairs and an array of variant T7 RNA polymerases to test, more insight into the DNA-protein interaction should be obtained during the course of the experiment.

Materials and Methods

Early Optimization

During early optimization of the experiment, the transcriptions were run with a 10x Ambion transcription buffer containing 400 mM Tris pH 7.8, 200 mM NaCl, 60 mM $MgCl_2$, 20 mM Spermidine HCl and 100 mM Dithiothreitol that requires 0.5 mM of each NTP. In addition, these reactions needed 0.5 U of YPP, 20 U of RNasin, and 5 U of T7 RNA polymerase. These transcriptions had low yields, and product formation plateaued at this very low concentration. These findings led to an experiment comparing the yield of the 10x Ambion transcription buffer and the 10x NEB high Mg transcription buffer. Results demonstrated that the NEB high Mg buffer had a higher yield but not at the desired level. The next optimization to solve this issue

dealt with the amount of T7 RNA polymerase used. A reaction was run using a ten-fold increase of T7 RNA polymerase concentration resulting in 50U of RNA polymerase in the reaction. This reaction increased product formation but also resulted in a smear of products on the gel, making it difficult to discern between product lines. Dilution of the 50U reactions aided in an increase in resolution; however, dilution still failed to yield the desired results.

Transcription and Gel Analysis

After the optimization, a 20µL reaction volume was used for T7 RNA polymerase reactions with wild-type promoters. The reaction components were 2 µL of a 10x New England Biolabs high-Magnesium reaction buffer and 0.8 µL of a 100mM solution of the 4 NTPs for a total of 3.2µL 100mM NTP mix, 0.5 U YPP, 20 U Rnasin, and 50U of the wild-type T7 RNA polymerase. The NEB high MG buffer contains 400 mM Tris pH 7.9, 200 mM $MgCl_2$, 100 mM Dithiothreitol, and 20 mM Spermidine. The buffer also requires a final NTP concentration of 4.0 mM each. In addition, 320 ng of each targeting and non-targeting DNA sequences containing the T7 promoter were used for template in the reaction. Prior to every transcription, 5 µL of each targeting and non-targeting DNA strand was mixed with 5 µL of 10x Ambion PCR buffer and 35 µL water, denatured at 80°C for five minutes and annealed at room temperature for five minutes. The transcription reaction was run in a standard Thermocycler holding at 37°C for 60 minutes. Sometimes 4 µL aliquots were taken at 0, 15, 30, and 60 minute intervals for analysis of product formation. Otherwise, a single 4µL was taken at the end of the 60 minute reaction. For each aliquot taken, 16µL of a 2x ML Stop buffer with EDTA and blue dye in addition to 20µL water was added to halt transcription. For gel analysis, a 20% SDS-PAGE gel was used. The contents of the gel were 25 mL acrylimide, 75 µL 10% APS, and 25 µL TEMED. 10 µL of the transcription reaction plus stop ML Stop dye mix were loaded into each well of the PAGE gel in

order of aliquot taken. These gels were run exposed to approximately 425 Volts for 2 to 2.5 hours. Gel runs were generally stopped when the stop dye reached an inch from the bottom of the gel plates. After gels were removed from the glass plates, they were mixed with 25 µL of SYBR Gold staining dye for thirty minutes. The SDS-PAGE gels were then imaged with a fluorescence scanner using 450 nm wavelength blue light to quantify production of RNA transcripts from the reaction.

Gel Analysis with Radiation

The previous work still failed to yield the desired resolution, so a switch was made to radiation for visualization. In addition to the normal NTPs, 1 µL of radioactive alpha-ATP was added to each reaction to be incorporated into the products. This yielded a much clearer picture on the gels; therefore, radiation was used for the duration of the experiment. When modified T7 RNA polymerases were used, the modified polymerase stocks were diluted by half to ensure similar concentrations of polymerases between the wild-type and modified reactions.

Experimental Design

To observe the interaction between the modified and wild-type T7 RNA polymerases with wild-type and unnatural base pairing, single base pair substitutions were made at the -9, -8, and -7 positions relative to the start site. These substitutions focused only on major groove interactions. Modified oligonucleotides were produced in outside labs and hybridized in an annealing reaction to produce double-stranded DNA for transcription templates. Overall, there were three alternative wild-type substitutions and four unnatural substitutions performed at each position. The four unnatural substitutions consisted of substituting isoC:isoG and dZ:dP in both orientations. After the 60 minute reaction, transcription levels were analyzed by way of SDS-PAGE gel analysis.

Oligonucleotide Synthesis

Wild-type and isoC/isoG DNA sequences were synthesized and HPLC purified by Integrated DNA Technologies (IDT). The DNA sequences containing dZ and dP were synthesized and gel purified by Zunyi Yang (Foundation for Applied Molecular Biology, Gainesville, Florida).

T7 RNA polymerase variant construction

Mutations to the T7 RNA polymerase gene were done using a megaprimer whole plasmid PCR[8]. These megaprimers were created by PCR using Invitrogen Platinum Taq High Fidelity. This round of variants only consisted of amino acid mutations at the 758 position of the T7 RNA polymerase. These variations were made using reverse primers containing the mutation needed to create these variants. 10ng of pQE80L wild-type T7 RNA Polymerase was cycled under PCR conditions 15 times with ED.T7 1342 F and the mutated reverse primers. The PCR reaction had a volume of 50 µL, an annealing temperature of 55º C, and an elongation time of 1 minute. PCR products were isolated and purified using Wizard Gel and PCR Clean-Up System by way of the Promega PCR Clean-Up protocol. Approximately 500 ng of the purified megaprimer was mixed with a 50 µL Pfu Ultra Hotstart PCR using Strategene products. 200 ng of pQE80L Wild-type T7 RNA Polymerase was run for 18 cycles with the cycles running at temperatures of 95º C for 30 seconds, 55º C for 30 seconds, and 68° C for 12 minutes. 20 U of DpnI made by New England Biolabs were added after the PCR and the total reaction was maintained at 37º C for one hour to digest the pQE80L Wild-type T7 RNA Polymerase template. DpnI was then heated to 80º C for 20 minutes for inactivation. 2 µL were transformed to OmniMax2 chemically competent cells, and cells were plated for culturing. Successful clones were isolated and purified. In addition, sequencing was performed to determine success.

T7 RNA polymerase variant expression and purification

T7 RNA polymerase was expressed in Strategene *E. coli* BL21 gold culture and grown in 2mL LB to the mid log growth phase. This culture was then diluted into 100 mL TB-Amp broth, which consisted of 47.6g Terrific Broth powder, 4mL glycerol, and 100mg of ampicillin per liter and grown at 37° C until an optical density of approximately 0.7-0.9 was observed. Next, the cultures were placed in ice for 30 minutes and then induced with 1 mM IPTG before being incubated overnight at 22 ° C. The following day, a cell pellet was formed by centrifuging the culture at 6,000 x 12 g for 10 minutes. Next, a 10 mL Binding Buffer consisting of 5 mM imidazole, 0.5 M NaCl, 50 mM pH 8.0 Tris buffer, and one pellet of Roche protease inhibitor was used for re-suspension. Then, the mixture was sonicated for 3 minutes at 50% strength. Cell debris was isolated by centrifuging the mixture at 9,000 x g for 30 minutes. Next, the lysate portion was first washed over 0.5 mL of Qiagen nickel resin, then washed with 10 mL of Binding Buffer, and then washed again with 4 mL of Wash Buffer (20 mM imidazole, 0.5 M NaCl, 50 mM Tris buffer pH 8.0) before being eluted in 4 mL Elution Buffer (250 mM imidazole, 0.5 M NaCl, 50 mM Tris buffer pH 8.0).

The Elution Buffer was replaced with Storage Buffer (50 mM Tris, pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) using Amicon Ultra-15 Centrifugal Filter Devices from Millipore. First, 5 mL of Storage Buffer was pipetted onto the filter and centrifuged at 3,250 x g for 15 minutes. Storage Buffer was added to the protein until the total volume was 15 mL. It was then pipetted onto the filter and centrifuged at 3,250 x g for 45 minutes. After this, 15 mL of Storage Buffer was mixed with the concentrated protein and placed in the centrifuge at the same speed for another 45 minutes. The final volume ranged from 0.2 to 1 mL with concentrations of approximately 10 mg/mL. This was diluted by half with 100% glycerol and stored at 20°C.

Protein levels were roughly quantified by absorbance at 280 nm, performing a range of dilutions to test, and then verified by SDS-PAGE (Pierce 4-20% Precise Protein Gels). The wild-type T7 RNA Polymerase and other variants were diluted to 1.0 absorbance at 280 nm.

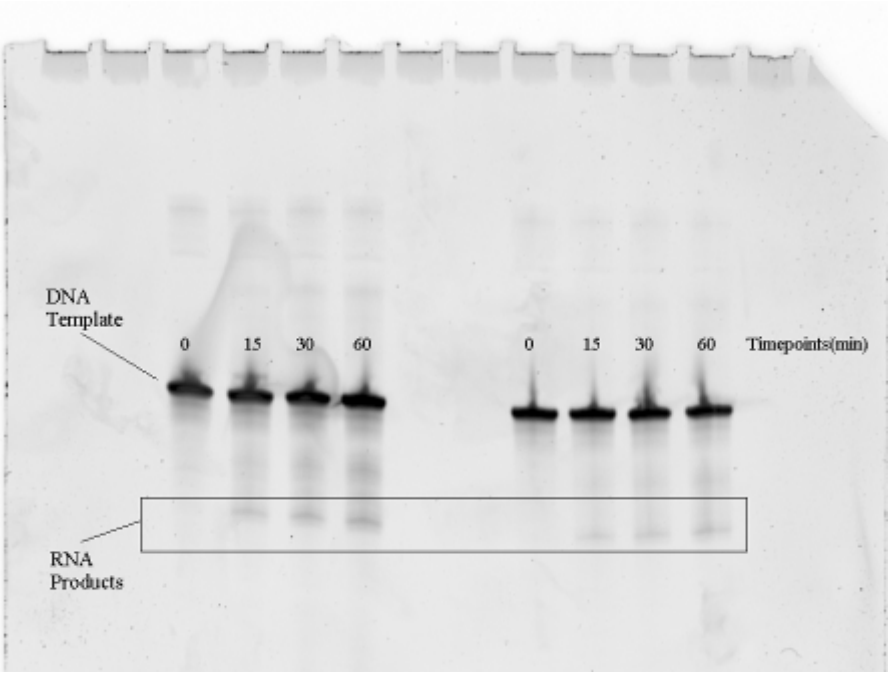
T7 RNA polymerase activity

After the purification step was complete, the isolated wild-type T7 RNA polymerase from the variant isolation procedure was tested alongside the commercial Stratagene T7 RNA Polymerase to estimate level of dilutions needed on the variant polymerases in order for them to perform well in the transcription reactions. Eight reactions were run, one with 1 µL of 50 Units Stratagene T7 RNA polymerase and seven reactions of serially diluted T7 RNA polymerase isolated from the other variants. The first lane had a half dilution of the original stock, with each successive lane being diluted by half. After analysis, the first dilution of the isolated T7 RNA polymerase was equivalent in product formation with the Stratagene T7 RNA polymerase. The remaining variants also were diluted by half during the course of the experiment.

Results and Discussion

Early on, the major focus of the experiment was to optimize the transcription reaction and the analytical techniques to be utilized throughout the course of the experiment. In the first step of the experiment, two wild-type T7 RNA polymerase reactions with wild-type promoter DNA were run in the initial 10X Ambion buffer reaction conditions. Samples were taken at 0, 15, 30, and 60 minute intervals, and the gel was stained with SYBR-Gold dye (**Figure 4**). The two major observations from this gel were that the DNA bands blew out the RNA product bands and that the RNA products never plateaued in production before the 60 minute mark. Because of these results, all future transcription reactions were run to a full hour each time . However, the DNA

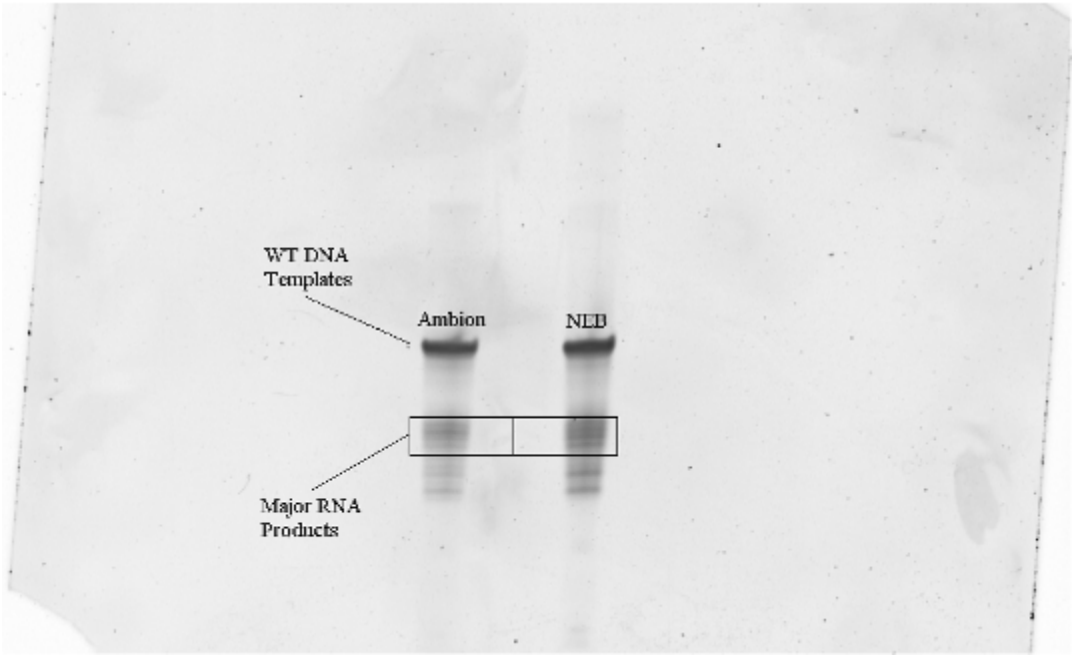
blowing out the RNA products signified a problem in the procedure. One explanation was that the DNA was annealing in the transcription reaction and in the process usurping some of the magnesium necessary for the reaction to run.



(**Figure 4**) Wild-type transcription test run: an 8% PAGE gel of two separate wild-type transcription reactions to test procedure; samples taken at various points in reaction and stained with SYBR-gold dye

A second transcription was run to test the effects of annealing the DNA prior to use in the transcription reaction as well as the effectiveness of the 10X Ambion reaction buffer to the New England Biolabs High Magnesiun reaction buffer. Before the transcription reaction, the targeting and non-targeting DNA strands were mixed in water and PCR buffer and then heated to 80°C for five minutes in order to anneal them prior to addition to the main reaction. The gel showed that the NEB high magnesium buffer and annealing of the DNA prior to the gel analysis resulted in a better production of RNA products (**Figure 5**). These also became standard steps performed

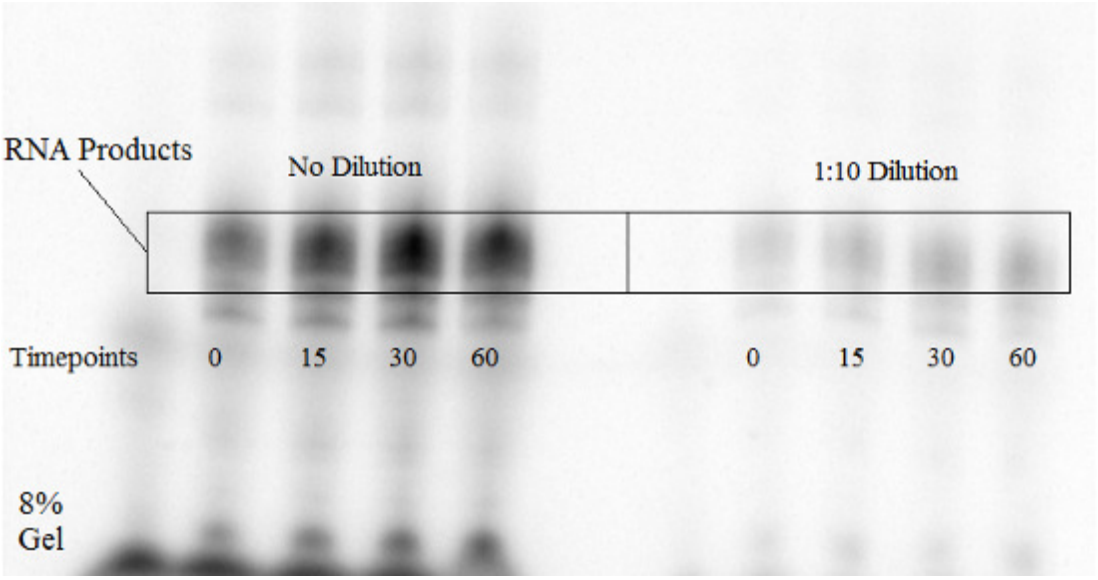
during the remainder of the experiment and seemed to consistently yield the best results during PAGE gel analysis. After this, further optimization was tested by varying the amount of T7 RNA polymerase present in the reaction and the density of the PAGE gel. The amount of T7 RNA polymerase was increased from 5 U to 50 U in each wild-type reaction, and the products were run on a 10% and 20% PAGE gel for analysis. These gels showed little improvement in the resolution, so radiation tagging was utilized instead.



(**Figure 5**) Ambion vs. NEB High Mg Buffer Comparison: two separate transcription reactions run with separate buffers for comparison. The gel is an 8% PAGE gel and SYBR-gold dye was used to visualized products.

For radiation, radioactive alpha-ATP was utilized. The first reaction run under radioactive conditions was on an 8% PAGE gel, and it showed significant improvement in resolution (**Figure 6**). The RNA products were clearly visible, and gel results showed that a 1:10

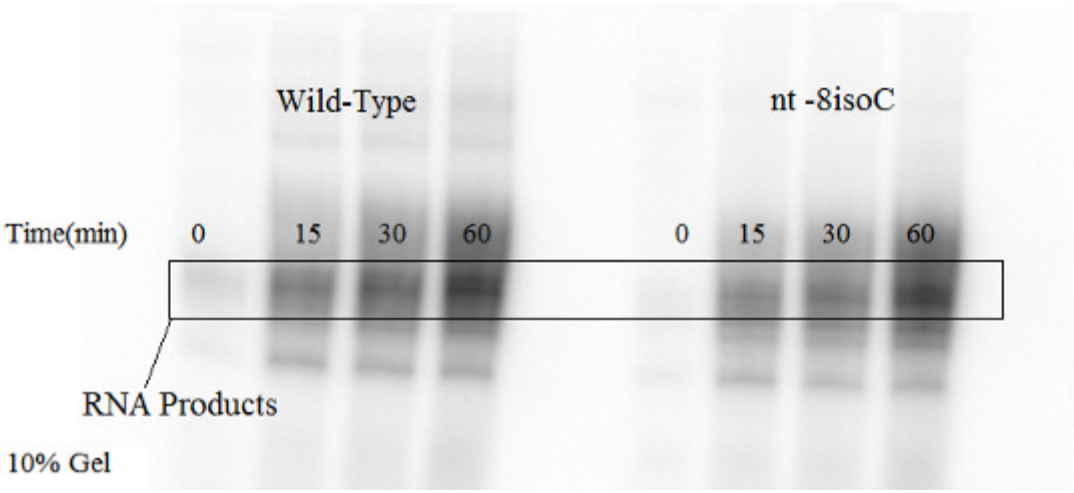
dilution of the products was no longer necessary before the gel run. However, resolution still was not at an optimum level, so further alterations were considered.



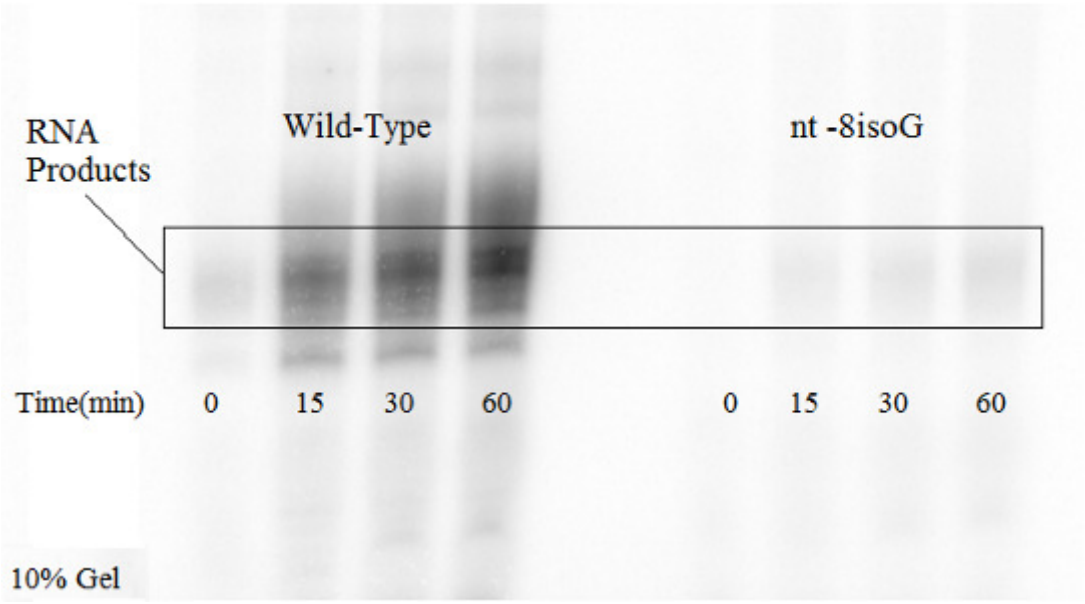
(**Figure 6**) First radioactive transcription: an 8% PAGE gel comparing two transcription reactions run with radioactive ATP. The second set is diluted by a factor of ten.

The next transcription reaction tested the first batch of modified base pairs, non-targeting -8 isoC and non-targeting -8 isoG. Non-targeting -8 isoC was expected to yield the same results as a transcription reaction using all wild-type components due to targeting -8 isoG's similar binding sites to the wild-type A base at that location. In addition, the process was modified by adding 36µL of stop dye to each 4µL aliquot taken, and they were run on a 10% gel. Gel results confirmed the hypothesized result of the reaction. The wild-type and non-targeting -8isoC were almost identical in quantity of products formed while non-targeting -8isoG showed little if any product formation (**Figures 7 and 8**). In addition, these findings also provided a verification of

the specificity of the T7 RNA polymerase for its promoter and gave indications that future substitutions would yield the same negative results.

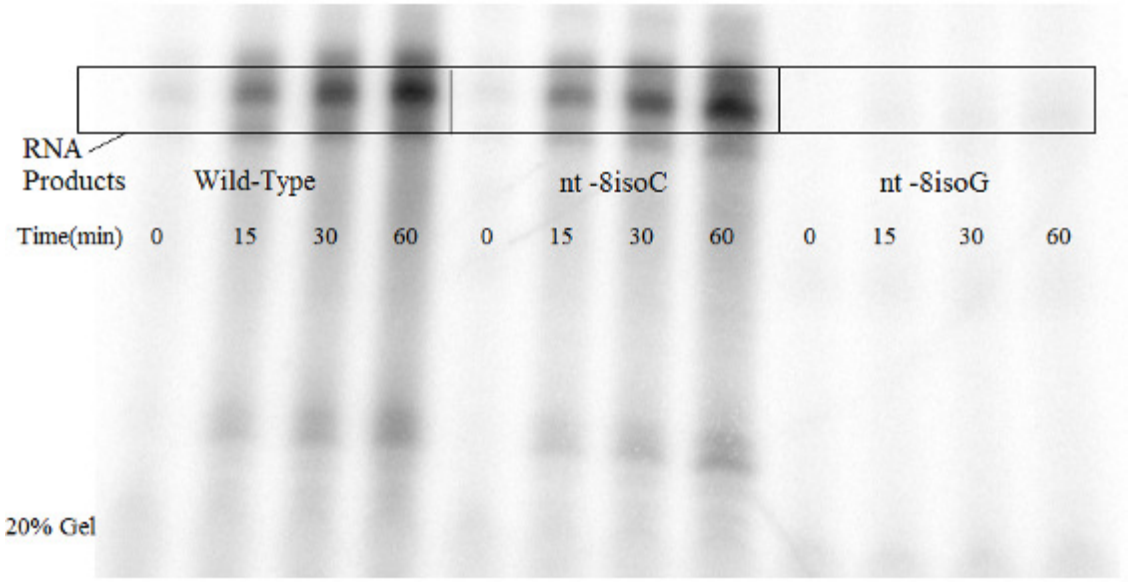


(**Figure 7**) Wild-type vs. nt -8isoC transcription: a 10% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and that of a modified promoter with an isoC substitution at the -8 position of the non-template strand



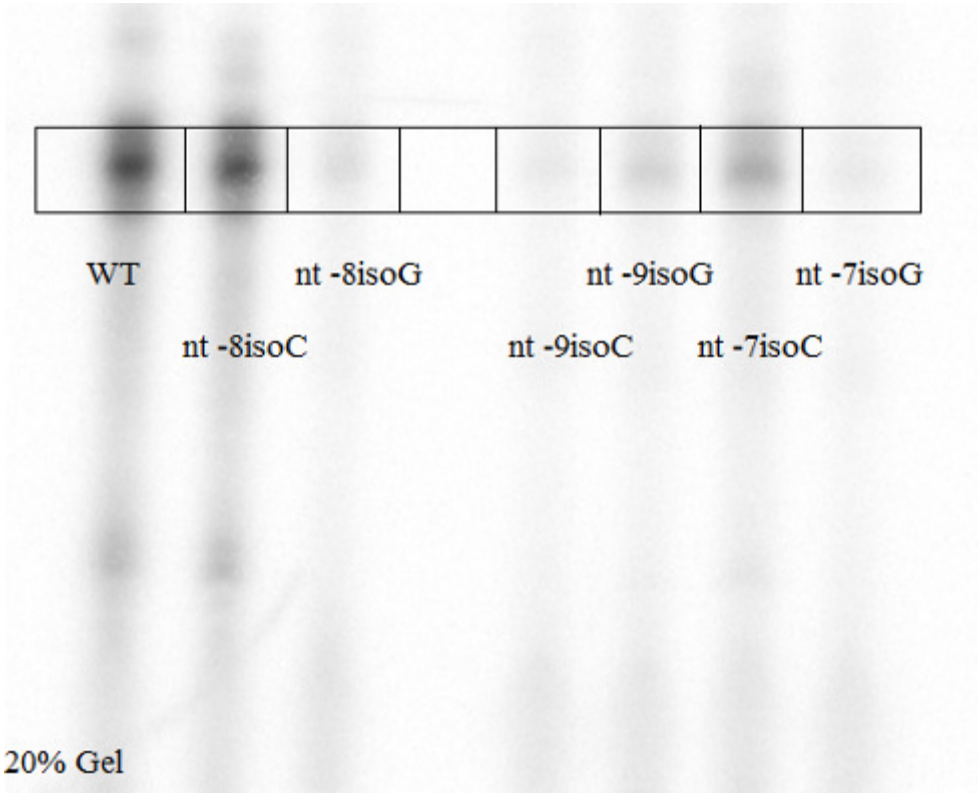
(**Figure 8**) Wild-type vs. nt -8isoG transcription: a 10% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and that of a modified promoter with an isoG substitution at the -8 position of the non-template strand

Resolution remained inadequate on the 10% gels, so the same process was performed with the exception of the stop dye and gel used. The previous reactions were repeated, and 16µL of stop dye plus 20µL of water were added to each 4µL aliquot. Furthermore, gels were analyzed on a 20% PAGE gel. This resulted in a much longer run time at 425V; however, the gel did not require the usual drying step. 20µL of each sample were loaded into the gel, and two separate gels were run to check for failure. Results showed a slightly better resolution where the major RNA products were grouped in a single area on the gel rather than the smear of products that had been occurring (**Figure 9**).



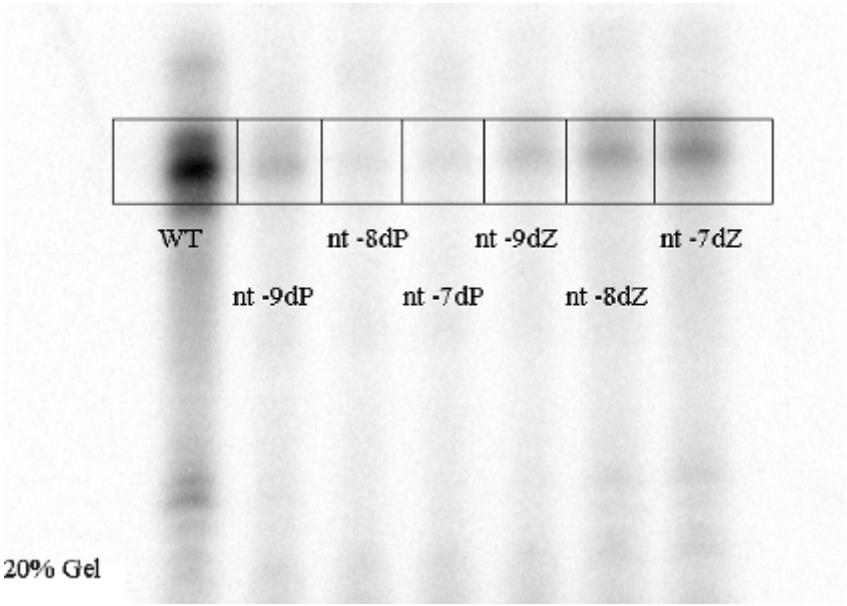
(**Figure 9**) Wild-type vs. nt -8isoG and nt -8isoC transcriptions: a 20% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and those of the modified promoters with each isoG:isoC substitution at the -8 position

The next step was testing the effects of replacing critical bases with artificial bases on transcription efficiency. This narrowed candidate bases to be used for the cross-testing with the modified T7 RNA polymerases. Efforts were concentrated on the -9 to -7 positions at first. There were a total of six substitutions analyzed during this round of transcription (**Figure 10**). As expected, the isoC substitution on the non-template strand at the -8 position resulted in an identical number of RNA products being produced at the same level as with the wild-type promoter. The remaining substitutions showed little if any transcription activity with the exception of the non-template iso Guanine substitution at the -9 position.

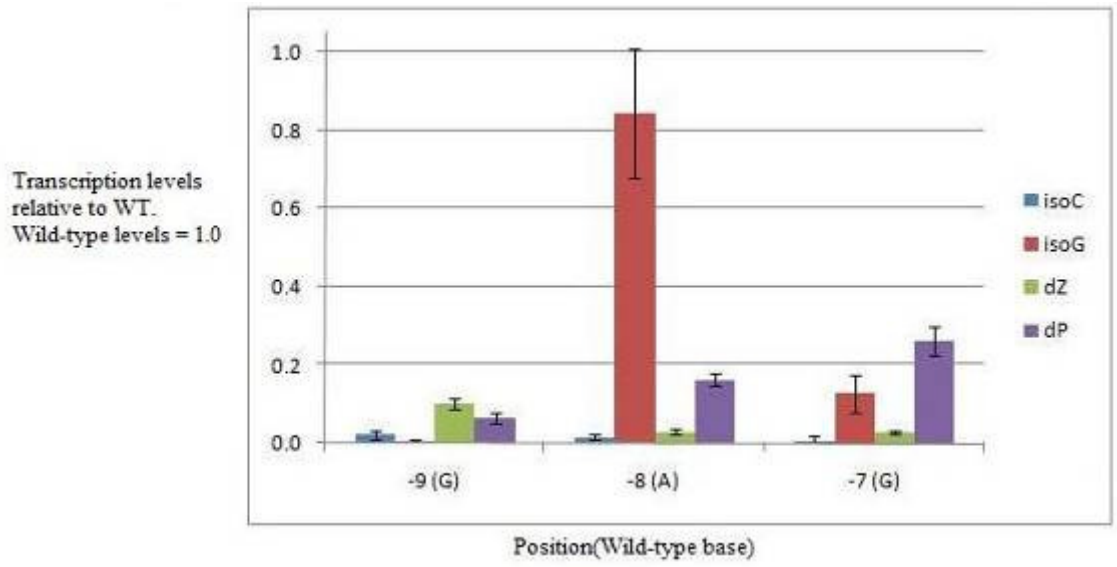


(**Figure 10**) Wild-type vs. isoC:isoG substitution transcriptions: a 20% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and those of the modified promoters with each isoG:isoC substitution at every position.

Furthermore, substitutions with the dZ:dP base pairs were carried out at the same positions. Results indicated the non-template -8 dZ and non-template -7 dZ showed some activity relative to the wild-type transcription, while the others showed minimal activity (**Figure 11**). Each of these reactions was triplicated, and relative activity was calculated and graphed. **Figure 12** illustrates the relative activity of each substitution versus the wild-type promoter. After compilation of the data was completed, there were some interesting items of note. As expected, the isoG substitution on the template strand at the -8 position showed little variance from the wild-type reaction. This was expected based on the fact that at this orientation, isoG:isoC resemble the wild-type A:T base pair. This phenomenon demonstrates that the most important factor in determining success of transcription is the major groove interactions between the function groups on the base and the T7 RNA polymerase. Base recognition does not seem to be a significant factor during the recognition process.

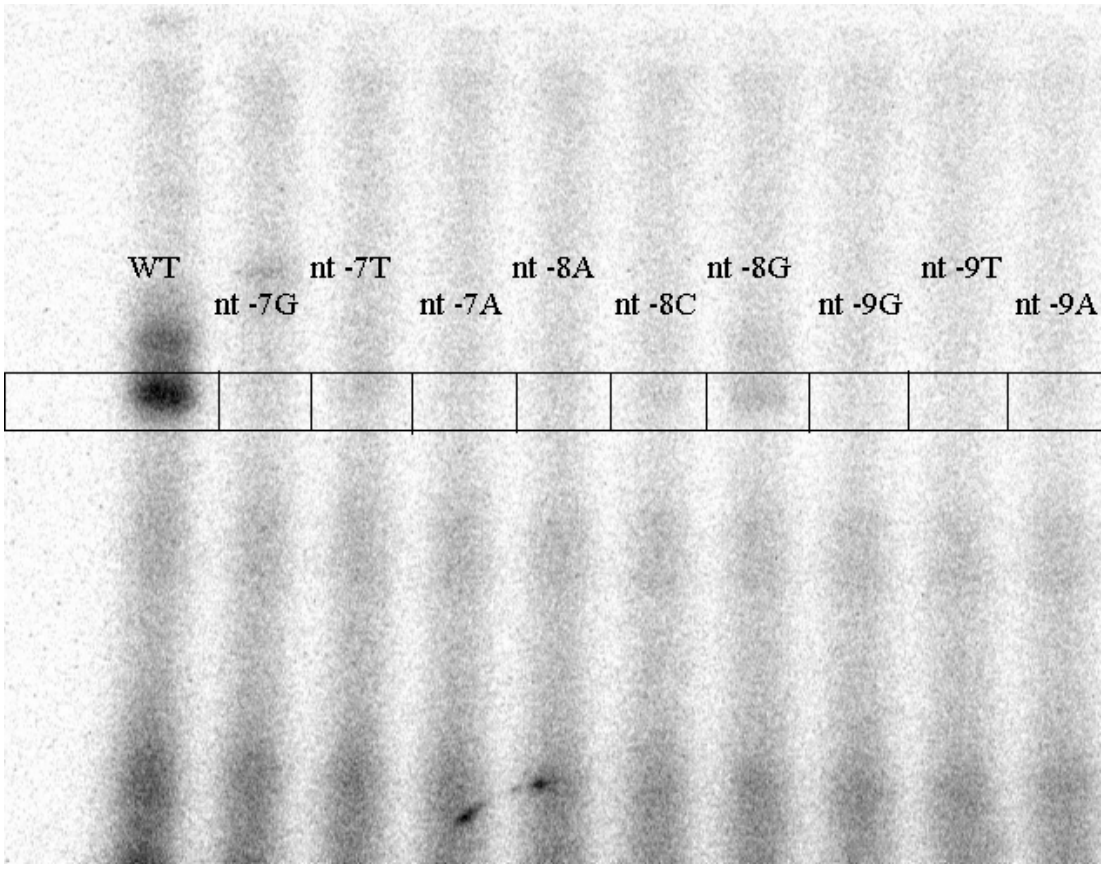


(**Figure 11**) Wild-type vs. dZ:dP substitution transcriptions: a 20% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and those of the modified promoters with each dZ:dP substitution at every position



(**Figure 12**) Relative transcriptional activity of variant promoters: compares the transcriptional activity of each modified promoter to the wild-type sequence. Wild-type transcription levels were normalized to 1.0; other transcription levels are graphed relative to wild-type.

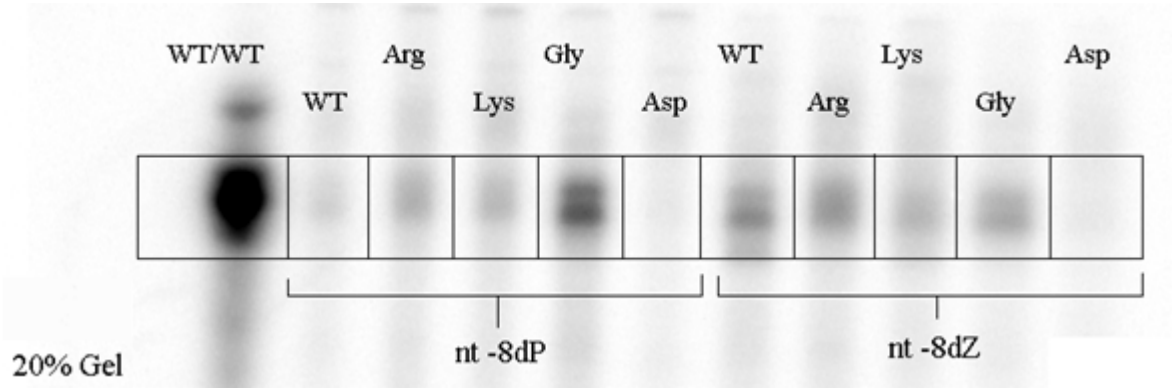
A final analysis on the sensitivity of the T7 RNA polymerase was performed using wild-type substitutions at the critical -9 to -7 positions. The three alternative wild-type bases were replaced at each position in order to confirm the effectiveness of substitutions in halting transcriptions. As the gel shows (**Figure 13**), all of the wild-type substitutions resulted in little transcription, proving once again that the major groove interactions between the T7 RNA polymerase and the functional groups of the base remain the most important factor for promoter specificity.



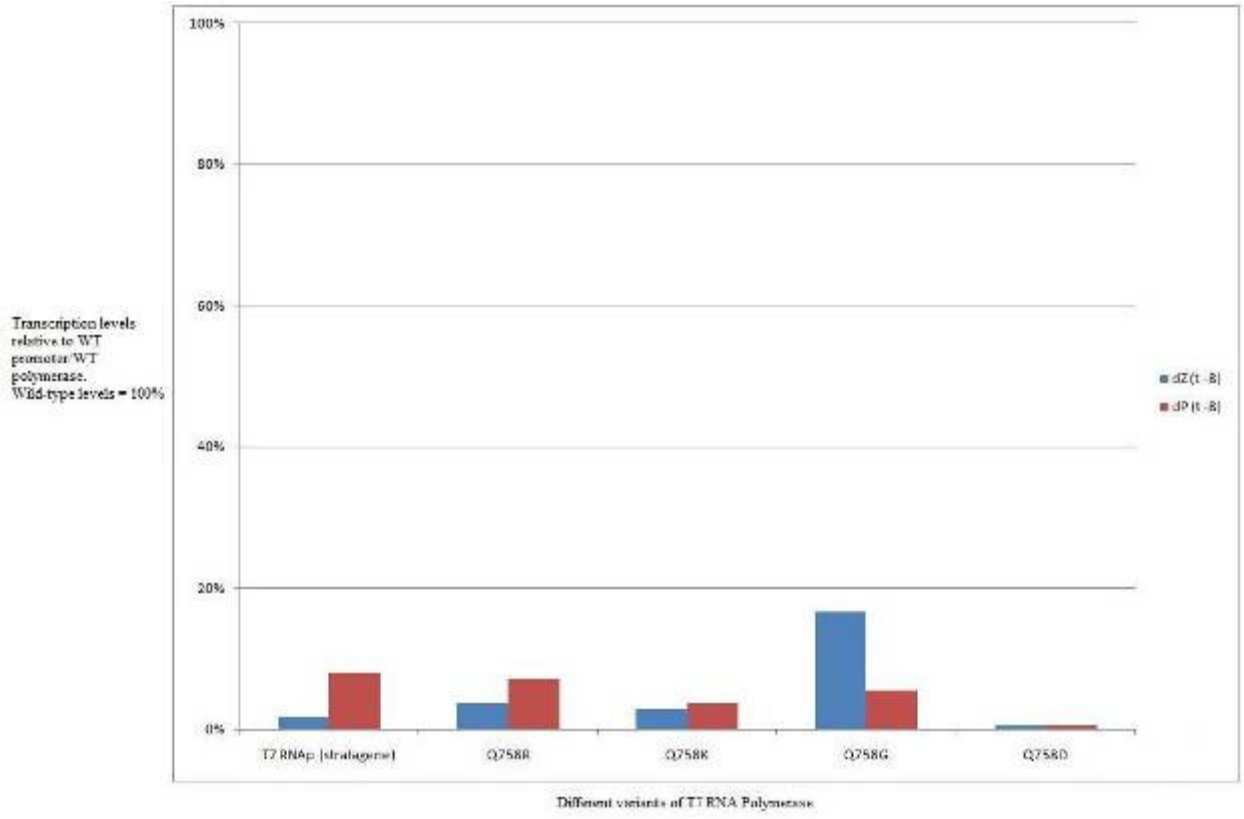
(**Figure 13**) Wild-type vs. natural base pair substitution transcriptions: a 20% PAGE gel using radioactive visualization that compares transcription product levels of a wild-type promoter reaction and those of the modified promoters with substitutions using natural base pairs at each position

From the T7 RNA Polymerase modifications, 4 separate strains were isolated during the time constraints of this experiment. All of the modifications occurred at the 758 position; normally a glutamine residue interacts with adenine at the -8 position in the major groove of the DNA. Of particular interest in these substitutions was testing the potential effects on the T7 RNA polymerase's interaction with the dZ:dP base pair substitutions in the promoter sequence. The four strains isolated included a 758-Arginine, 758-Lysine, 758-Glycine, and a 758-Asparagine

variant. In addition, as additional control, a wild-type T7 RNA Polymerase strain was isolated using the same procedure to be tested alongside the other variants. Since all of the variations occurred at the 758 position of the T7 RNA Polymerase, the first transcription test was performed on the two dZ:dP substitutions at the -8 position. As shown before, these substitutions were strong candidates due to their highly negative effects on transcription levels when coupled with the wild-type T7 RNA polymerase. When exposed to the wild-type T7 RNA polymerase, the template dZ substitution yielded less than three percent of the normal transcription levels while the corresponding dP substitution yielded under seventeen percent of the normal transcription levels. Transcription with the dP:dZ substitutions and all four T7 RNA Polymerase variants was performed, and the gel yielded promising results (**Figures 14 and 15**). Of particular interest was the template -8 dZ substitution. The wild-type commercial T7 RNA polymerase had little transcription activity, while three of the modified polymerases showed more activity with this dZ substitution. In particular, the 758-glycine mutant indicated ten times the transcriptional activity as the commercial wild-type T7 RNA polymerase. The success of this transcription reaction demonstrates that unnatural base pairs possess the ability to take a DNA sequence to RNA successfully through transcription. With more time in the future, numerous variations of this experiment may be performed to perfect the binding affinity of variant T7 RNA polymerases to promoter sequences that contain abnormal base pairing.



(Figure 14) Comparison of T7 RNA Polymerase Variants with -8 dZ:dP substitutions: a 20% PAGE gel using radioactive visualization that shows transcription levels produced by each WT and variant T7 RNA polymerase when exposed to the two -8 dZ:dP substitutions



(**Figure 15**) Relative transcriptional activity of variant T7 RNA polymerases: compares the transcriptional activity of each polymerase with a dZ:dP modified promoter at the -8 position to that of a wild-type promoter with a wild-type polymerase. Wild-type transcription levels were normalized to 100%; other transcription levels are graphed relative to wild-type.

Problems Encountered

While early optimizations were frustrating because of the lack of consistency and quality in the gel images, eventually the radiation protocol led to gel images that yielded clearer results for calculations. Still, gel images show a multitude of variation in quality and clarity of the images. One solution that improved this early on was changing to 20% SDS-PAGE gels which gave a higher resolution due to the slow pace of RNA migration. Also yielding positive results was the heat-denaturing of the RNA-stop dye mix prior to addition to the gel. Both of these factors improved the quality of the images and band resolution, yet they failed to provide the resolution desired. Possible solutions to remedy this may be loading less material per well or degrading DNA in the mixture to decrease any interference the DNA may be causing during the RNA migration.

In addition, there were problems during an attempt to form a radioactive RNA size marker by way of a kinase reaction. At each step of the procedure, a Geiger Counter was used to check for radioactivity, and multiple times the purified eluate that filtered through the column had zero radioactivity while the column itself continued to give high radioactive readings. Multiple troubleshooting measures were implemented, including replacing the kinase and buffer during subsequent reactions; however, none of these measures solved the problem. The columns were rated for 25 base pair or longer RNA; therefore, this should not have been the issue since the RNA molecules were 30 base pairs in length. Nonetheless, when the column purification step was skipped and the contents of the eluate were run on a gel, clear RNA bands appeared at the correct location on the gel. These bands travelled the same distance as the products of the +19wt reaction done earlier, which also were approximately 30 base pairs in length. Therefore, the columns seemed to be binding the RNA molecules during centrifugation and preventing flow-

through. Even without the column purification, the newly radioactive RNA markers continued to run in the correct fashion expected.

Conclusion

Overall the experiment was quite successful and provided the groundwork for further study in a new field. The isoC:isoG experiments at the -8 position demonstrated that correct hydrogen bonding between the 758 Glutamine group was sufficient to produce successful transcription even though isoC:isoG remain fundamentally different molecules from adenine and thymine. This successful mimicking demonstrated that the T7 RNA polymerase's ability to initiate transcription around the promoter site was mostly driven by major groove interactions rather than base recognition. Therefore, even though the T7 RNA polymerase has a high specificity for its promoter, base recognition has little to do with this characteristic. Another interesting project may entail studying promoter properties at the -1 to -4 and -13 to -17 positions to determine the effect of unnatural base pair substitutions on polymerase binding. Unlike the previous experiments, the isoC:isoG base pairs do not mimick an A:T base pair when interacting with the T7 RNA polymerase from the minor groove. Results in the major groove experiments suggest the probability that minor groove interactions rather than base recognition may also be the major determining factor for the success of transcription.

In addition, several interesting dZ:dP interactions were observed during the course of the experiment. While the isoC:isoG experiments mostly dealt with mimicking natural base pair interactions with the T7 RNA polymerase, the dZ:dP substitutions introduced unnatural functional groups such as nitrogen dioxide for major groove interactions. At the -8 position on the template strand, the dZ base led to minimal transcription levels when using the wild-type T7

RNA polymerase; however, it led to a ten-fold increase in transcription when the 758 residue was replaced with glycine. Further successes such as these increased transcription levels may have implications for future research. Incorporating unnatural base pairs into natural biological systems could lead to an increase in the variability of genetic code in determining transcription levels. In addition, unnatural amino acids could be incorporated into the T7 RNA polymerase for interactions with the modified promoters. This would lead to a better understanding of protein-DNA interactions and production of proteins tolerant to unnatural base pairs. Overall, incorporation of these unnatural elements may expand the variety of genetic code and amino acid components of proteins and offer new and interesting structures to study and apply to biological research.

Bibliography

1. Cheetham, G.M.T., et al. (1999) "Structural basis for initiation of transcription from an RNA polymerase-promoter complex." Nature **399**: 80-83.

2. Jung, K.H., et al. (2005) "Nucleotide analogues as probes for DNA polymerases." Cell. Mol. Life Sci **62**: 2080-2091.

3. Fukunaga, R., et al. (2008) "Phosphoserine aminoacylation of tRNA bearing an unnatural base anticodon." Biochem Biophys Res Commun **372**(3): 480-485.

4. Yang, Z., et al. (2007) "Enzymatic incorporation of a third nucleobase pair." Nucleic Acids Research **35**(13): 4238-4249.

5. Watson, J.D. and Crick, F.H.C. (1953) "Molecular structure of nucleic acids." Nature **171**: 737–738.

6. Gao, J., et al. (2004) "Expanded-size bases in naturally sized DNA: evaluation of steric effects in Watson-Crick pairing." J Am Chem Soc **126**(38): 11826-11831.

7. Hirao, I., et al. (2006) "An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA." Nat Methods **3**(9): 729-735.

8. Miyazaki, K., et al. (2003) "Creating random mutagenesis libraries by megaprimer PCR of whole plasmid (MEGAWHOP)." Methods of Molecular Biology **231**: 23-28.